May 17, 2019

Forensic Statistics, Online Course

INSTRUCTOR: Distinguished Professor Cliff Spiegelman

Office 459A Blocker, phone 979-705-6437

E-mail cliff@stat.tamu.edu (or please call 979-705-6437 and, if I do not pick up, leave a message including your phone number. I will call you back.)

TEXT: Ott and Longnecker: Statistical Methods and Data Analysis (7^{th} edition).
Additional reading material:
https://www.sfu.ca/~palys/Campbell&Stanley-1959-Exptl&QuasiExptlDesignsForResearch.pdf

Reference Manual on Scientific Evidence: Third Edition

Also Donoghue: Statistics and the Media; Foundations in Statistical Thinking through Media

Required: JMP (software provided).

Course description and prerequisites:

Forensic statistics is a continuing ed stat course designed for forensic practitioners. You will have exposure to a broad range of forensic examples including those from GSR, firearm/toolmarks, DNA, glass, bomb fragments, fingerprints etc. You are expected to work hard to understand the context and lessons from all examples whether or not they come from your specialty. In addition, I rephrase a quote from President J. F. Kennedy 'Ask not what Professor Spiegelman will do to teach you statistics, but rather how hard you will work to learn statistics.' I will my do my part to help you work to learn the material for this course.

We will cover approximately the first 60% of the text, that is, chapters 1-11. Please read the relevant part of the text and references, particularly Reference Manual on Scientific Evidence as we go along. We will cover most of the methods and ideas in these sections of the text. We will also present material, with emphasis, that is not in the text such as database search false match probabilities, and equivalence tests, and the scientific method.

Course Objective:

1. Students should understand the need for well-designed experiments and surveys.
2. They should be able to assess the designs used in published papers to make claims of scientific advance. Students will be asked to evaluate designs and their strengths and weaknesses.
3. They should be able to read the textbook and similar texts to update their knowledge as statistical practice changes. They should also understand the limits of their statistical knowledge, so as to keep them out of professional disrepute. It is also expected that you will be able to identify situations (that may occur in your career) where you will either need to ask for additional professional statistical help or need take a more advanced course.
4. They should be able to calculate needed sample sizes for basic experimental designs and surveys.

Accessing the Course

1. Enroll in the course using the link below:

   https://statistics.catalog.instructure.com/courses/forensic-stat
2. After registration and payment have been received, you will be given the link and login information to access the learning management system used for the course.
3. The course site will include several modules including the lecture materials, datasets, and supplemental lectures. A previous exam will be found in the modules area as well.
4. You will have 6 months of access to the course.

Software Information

Any demonstrations are done using JMP software. Once you are registered for the course, you will receive instructions and a link to download the software.

Exam Information

Final exams will be given in person in Austin, TX three times a year. The first available exam date for which registered students who have completed the course may take the exam is August 2, 2019. Future exam dates will be announced as that information becomes available. Before each exam, a review session will be held.

- Lecture 1: Welcome, Syllabus, and course culture, What Is Statistics and Why Is Statistics Important? Observational Studies and Surveys, What Is An Experiment, Repeatability and Reproducibility, Accuracy and Precision, Internal, External, And General Validity, Pre-Experimental Designs, A True Experimental Design, Design Evaluation, Internal Validity, External Validity,
- Lecture 2: Review Lecture 1 plus useful plots: Pie Charts, Histograms, Boxplots, q-q plots, Scatterplots, Scatterplot Matrix, Parallel Plots, Time Series Plot, and Control Charts
- Lecture 3: Review Lecture 2 Fallacies Of Presumption, Independence: An Abused Assumption , Randomization: The Wonder Drug For Avoiding Bias And Obtaining Representative Samples , convenience samples , Randomizing And Using Sufficient Sample Sizes Will Greatly Help Achieve External Validity, Hypotheses Need To Be Properly Stated, Continuous Distributions, Normal (the most common, but not always appropriate, continuous distribution, & used for testing, confidence intervals), t (used for testing, confidence intervals, & modeling data with many outliers), Chi-Square F Distribution, Standard normal distribution , How to transform any normal distribution to a standard normal distribution and vice versa, F-Distribution, Discrete Distributions, Binomial distribution, Poisson distribution and pmf, Negative Binomial distribution, Hypergeometric Distribution
- Lecture 4: Review of Lecture 3 plus Probability: The Language Of Chance, Interpretation of probability, Three dialects (flavors), Sample space and events, Set Theory, Probability rules, Key concept: Independence, Independence is important but overly used. , Cumulative distribution functions and probability density functions, Probability mass function, Expectations of random variables, The limiting value of the sample mean is the population mean; , Means behave properly for unit changes such as km/h to mi/h, Variances, Covariances, and correlation of random variables, Propagation of error , A couple of additional propagation of error formulas, Conditional Probability, Bayes Theorem, Likelihood Ratios + Bayes' Factors
- Lecture 5: Review of Lecture 4 plus Audit Sampling Of Questionable Work Products, Starting place: Why Hasn't This Been Done Before?, We've had these problems before, but why haven't we had a framework in place already?, Sampling, Fair and meaningful collection of data is a challenge for forensic disciplines., Most practices were developed with little input from statisticians, Often Heard Remarks: We live in the real world:, We have a representative

database of bullet measurements: Famous last words of CBLA, Starting Place, The Sampling Frame, Sampling Frame:, Subpopulations?, What A Sample Of Size 2 Looks Like For N=4: SWR & Order Does Not Matter, What A Sample Of Size 2 Looks Like For N=4: SWOR & Order Does Not Matter, Why Use The Hypergeometric Distribution?, Examples

- Lecture 6: Review of Lecture 5 plus Hypotheses testing framework, What is the Population, The Null Hypothesis for a Single Mean, The Alternative Hypothesis, Choosing the null hypothesis, The Choice Of Null Hypothesis Has Consequences On The Outcome, Type I Error (False Reject), Type I Error Rates and Confidence level, Choose a Decision Rule, All (simple) hypothesis are wrong but some are useful. (Sound familiar?), Type II: The Other Kind of Error, Statisticians spend a lot of time trying to figure out a priori if a study is large enough to detect meaningful departures from a null hypothesis, Z-Test Examples, P-values, Examples, Problems with Significance Tests, Logic Behind Significance Tests, Lesson: The chance for both hypotheses needs to be calculated., Likelihood Ratios, Example: Retroactive Case Reviews
- Lecture 7: Review of Lecture 6 plus Inference about two population means, Notation, The 3 t-tests that are the work horse of two population comparisons, Paired t-test, Pooled t-test, Unequal variance or Welsh t-test, Pooled t-test, Examples and plots, t-Test Template, Examples, Cautions using case with hypotheses about counts, Equivalence t-Tests, Examples, Interval estimation, Example: Refractive Index of Glass by Type, Example: Confidence Interval For Headlamps By Hand, JMP CI demo, Effect of Sample Size, Effect Of Estimated Standard Deviation, Using JMP to Construct CI, Sample Size Determination, Tests and confidence intervals for a single proportion, We require a minimum of 5 successes (1s) and 5 failures (0s), The Confidence Interval, The Special Case Of No Successes Or No Failures, Second Caveat, A better approach, Justification for adding 2 successes and 2 failures, Exact tests for small sample sizes, The setup, The test for equal proportions, Confidence interval for the difference of proportions, Caveats and adjustments are similar to one proportion case, Adjusted intervals for the difference of proportions, Poisson Distribution, Example: GSR For Non-shooters, Likelihood Ratios
- Lecture 8: Review of Lecture 7 plus Examples Where The Binomial Distribution Is Used Whether It Should Be Used Or Not

- Lecture 9: Review of Lecture 8 plus Analysis Of Variance, Making inferences when there are 3 or more populations, This is called ANOVA, It is somewhat more formula dense than what we have been used to, Tests for normality are also somewhat more complex, ANOVA, We are going to consider a simple form, namely comparing 3 or more populations., Motivating Examples, There is considerable controversy as to how to compare 3 or more populations. Approaches., In general, if there are t populations, there are t(t-1)/2 two-sample comparisons., The controversy revolves around the concept of Type I error, Specifically, if we do 6 different 95% confidence intervals, what is the probability that one or more of them do not include the true mean?, It is higher than 5%!, If you do 20 confidence intervals, you expect 1 = 20 x 5% will not include the true population parameter, The first step is the ANOVA F-test, The F-test is easy computed, and provided in all statistical packages, Interpretation, Notation
- Lecture 10: Review of Lecture 9 plus Multiple comparisons, controlling experiment wise error rates or false discovery rate, Residuals as a means of checking the normality assumption
- Lecture 11: Review of Lecture 10 plus Chi-Squared tests for Independence and Homogeneity, Motivation, Null and alternative hypotheses, The Chi-square test for independence of two categorical factors, Using JMP, Firearm/Toolmark Example, Interpreting Output, Examples
- Lecture 12: Review of Lecture 11 plus Regression, Terminology and Notation, Motivation, Scatterplots, Examples, Least Squares Method, Interpretations of least squares results, JMP demo , Population parameters: , Sample statistics: , Underlying Assumptions
- Lecture 13: Review of Lecture 12 plus Are Y and X related, Inference about a population slope, Residual plots to test for normality
- Lecture 14: Review of Lecture 13 plus Finish correlation, More on outliers, More on leverage points

University regulations:

Americans with Disabilities Act (ADA) Policy Statement

- - The Americans with Disabilities Act (ADA) is a federal anti-discrimination statute that provides comprehensive civil rights protection for persons with disabilities. Among other things, this legislation requires that all students with disabilities be guaranteed a learning environment that provides for reasonable accommodation of their disabilities. If you believe you have a disability requiring an accommodation, please contact Leigh Savage at Leigh.Savage@fsc.texas.gov.
- Academic Integrity Statement and Policy
- - "An Aggie does not lie, cheat or steal, or tolerate those who do." http://aggiehonor.tamu.edu
- COPYRIGHT NOTICE: The handouts used in this course are copyrighted. By "handouts," I mean all materials generated for this class, which include but are not limited to syllabi, exams, lab problems, in-class materials, review sheets, additional problem sets, and recordings. Because these materials are copyrighted, you do not have the right to copy the handouts or videos, unless I expressly grant permission. You are allowed (expected) to print or watch the materials for your own use. I have given the Pre-Vet Society Test Bank permission to post exams from this course provided they are available to all TAMU students.
  - STATEMENT ON PLAGIARISM: As commonly defined, plagiarism consists of passing off as one's own ideas, words, writing, etc., which belong to another. In accordance with this definition, you are committing plagiarism if you copy the work of another person and turn it in as your own, even if you should have the permission of that person. Plagiarism is one of the worst academic sins, for the plagiarist destroys the trust among colleagues without which research cannot be safely communicated. If you have any questions regarding plagiarism, please consult the latest issue of the Texas A&M University Student Rules, under the section" Scholastic Dishonesty."
  - Academic Integrity Statement: "An Aggie does not lie, cheat, or steal or tolerate those who do."

The Aggie Honor Council Rules and Procedures are available at the web site: http://aggiehonor.tamu.edu.

See: http://ucc.tamu.edu/syllabus.html for more details.